Journal of Experimental Child Psychology 176 (2018) 55–72

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Building a knowledge base: Predicting self-derivation through integration in 6- to 10-year-olds

Alena G. Esposito *, Patricia J. Bauer

Department of Psychology, Emory University, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 16 January 2018
Revised 10 July 2018
Available online 10 August 2018

Keywords:
Self-derivation
Semantic knowledge
Integration
Verbal comprehension
Academic achievement
Transfer

ABSTRACT

Self-derivation of new factual knowledge through integration of separate episodes of learning is one means by which children build knowledge. Content generated in this manner becomes incorporated into the knowledge base and is retained over time; successful self-derivation predicts academic achievement. Yet the component processes on which self-derivation through integration depends are as yet unknown. In parallel studies with 6- and 8-year-olds ($N = 41$; Experiment 1) and 8- and 10-year-olds ($N = 40$; Experiment 2), we tested a number of predictors related to other productive processes and learning (reasoning, executive functions, verbal comprehension, and long-term retrieval). Across studies, with different methods, only verbal comprehension, a measure of accumulated semantic knowledge, accounted for unique variance in self-derivation through integration performance. The results indicate that self-derivation through integration of separate episodes relates to accumulation of knowledge and the ability to recruit the knowledge in the service of specific task demands. Implications for cognitive training and transfer are discussed.

* Corresponding author.
E-mail address: aesposito@clarku.edu (A.G. Esposito).

https://doi.org/10.1016/j.jecp.2018.07.011

Introduction

Constructing a knowledge base is one of the most important tasks in development. Knowledge lays a foundation for language and thought, and knowledge supports the acquisition of more knowledge (Goswami, 2011). Children build knowledge through direct tuition and experiences. Because learning experiences are distributed over time and modalities, for knowledge to accumulate, separate but related episodes must be integrated with one another. Importantly, elaborating the knowledge base also depends on productive processes that permit self-derivation of new semantic content based on integrated representations (see Bauer, 2009, 2012). Research has revealed that new facts resulting from integration and self-derivation become incorporated into the knowledge base (Bauer & Jackson, 2015) and are retained over time (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016). Critically, performance on tasks that test self-derivation of new factual knowledge through integration predicts academic achievement in both reading and math (Esposito & Bauer, 2017). Currently unknown are the component cognitive abilities on which self-derivation through integration depends. In the current research, in separate studies with 6- and 8-year-olds (Experiment 1) and 8- and 10-year-olds (Experiment 2), we tested candidate predictors suggested by research on other productive processes, specifically reasoning, executive functions, verbal comprehension, and long-term retrieval.

Productive processes and component cognitive abilities

Productive processes such as analogy, deduction, and induction allow learning to proceed efficiently because learners can extend beyond information explicitly provided in a lesson to derive novel understandings (e.g., Goswami, 1992, 2011; Perret, 2015). Productive processes have been observed across the lifespan (see Mandler & McDonough, 1996, for examples even during infancy) and are assumed to be a major mechanism of cognitive development (Bauer & Varga, 2017; Bauer, 2012; Brown, Bransford, Ferrara, & Campione, 1982; Goswami, 2011; Siegler, 1989).

The current literature on the component cognitive abilities on which productive processes depend has examined only a limited number of candidates, with particular emphasis on inhibitory control and working memory. The number of direct tests of these candidate processes is small, yet there is consensus that both inhibitory control and working memory play a role (see Goswami, 2011, for discussion). In terms of direct tests, Moutier (2000b) studied inhibitory control, training in logic, and deductive reasoning in 10- and 11-year-old children. In one condition children were trained to inhibit a common perceptual matching error, and in another condition they received logical training. At posttest, only the children with the training in inhibitory control showed improvement in the deductive reasoning task. In a related study, Moutier (2000a) tested 10- and 11-year-olds' deductive reasoning and performance on a number of inhibitory control tasks (Stroop Task, Trail Making, and Wisconsin Card Sort). Higher performance on the deductive reasoning task was related to higher inhibitory performance (see also Moutier, Plagne-Cayeux, Melot, and Houdé, 2006, for additional, indirect evidence of relations between cognitive inhibition and 8- to 10-year-olds' performance on deductive reasoning problems).

Other studies have not measured inhibitory control directly but have implicated it, working memory, or both based on patterns of task performance. For example, Richland, Morrison, and Holyoak (2006) investigated analogical problem solving in children aged 3–14 years. They found that younger children were more susceptible to the presence of feature-level distractors, relative to older children, and had lower levels of performance. They interpreted the pattern to indicate that younger children did not yet have sufficient inhibitory control to ignore the distractors and maintain the task-relevant information in working memory.

In addition to inhibitory control and working memory, accumulation and use of prior knowledge has been implicated as a factor accounting for age-related changes in inductive and deductive reasoning (e.g., Goswami, 2011). Yet this conjecture has not been subjected to strong empirical testing because there are few studies that have included a measure of verbal comprehension or other measures of acquired knowledge among the candidate predictor variables. An exception is Fisher, Godwin, and Matlen (2015), in which measures of verbal comprehension and reasoning were exam-

ined along with working memory and inhibitory control as predictors of inductive generalization with familiar categories. Yet in this study the verbal comprehension and reasoning measures were combined into a general intelligence composite, thereby precluding examination of their unique contributions to performance.

Another literature that bears on the question of relations between accumulation and use of prior knowledge and productive processes is research on inference making in the context of reading comprehension. Reading comprehension is strongly correlated to verbal comprehension (e.g., Riedel, 2007; Yovanoff, Duesbery, Alonzo, & Tindal, 2005), and both are described as measures of semantic knowledge and its recruitment and use (e.g., Akshoomoff et al., 2013; Schrank, 2010). Consistent with suggestions of relations between prior knowledge and successful inference, several studies have found correlations between reading comprehension and inference making within a single text such as drawing an inference from two consecutive sentences in a story (e.g., Cain & Oakhill, 1999; Cain, Oakhill, & Bryant, 2004; Oakhill, Cain, & Bryant, 2003). Interestingly, in this literature the question is whether inference making predicts reading comprehension, not whether reading comprehension—as an index of accumulation and use of prior knowledge—predicts inferential ability. This orientation is apparent in Cain et al. (2004), for example, where it was found that both working memory and inference making predicted unique variance in reading comprehension in children aged 7–9 years. This and other such studies suggest that there are relations between inference making and accumulated knowledge, yet they do not directly inform relations between acquired knowledge and inference as a productive process.

In summary, productive processes such as induction and deduction are supported by inhibitory control and working memory. Reading comprehension is related to these component cognitive abilities as well as inference making, suggesting that accumulated knowledge may also be related to productive processes. Thus, both reasoning and verbal comprehension are indicated in productive processing but have not been directly tested, and it is not clear whether they both explain unique variance.

Self-derivation through integration and component cognitive abilities

The specific productive process that is the subject of the current research is self-derivation of new factual knowledge through integration of separate yet related episodes of new learning. This process is of particular significance because it is a viable model of learning as it occurs outside the control of the laboratory. Initial investigations of this specific productive process used a story passage paradigm in which children were presented with true but previously unknown facts (e.g., dolphins live in groups called pods) embedded within richly contextualized story passages. After a brief delay, a related true but previously unknown fact (e.g., dolphins talk by clicking and squeaking) was presented via a separate story passage that had a unique story line. After another brief delay, children were asked questions (e.g., how do pods talk?), the answers to which could be generated by integrating the information from the separate episodes (e.g., Bauer & San Souci, 2010). Performance on control trials (in which only one of the two stories is presented) makes clear that both facts are necessary to answer the questions (Bauer & Larkina, 2017; Bauer & San Souci, 2010).

Consistent with the suggestion that self-derivation through integration is important to learning, new factual knowledge derived through this process is retained over time. For example, in Varga and Bauer (2013), 6-year-olds correctly produced the novel facts on 63% of open-ended trials and recalled 60% of the facts 1 week later (see Varga et al., 2016, and Varga & Bauer, 2017, for similar findings with 4-year-olds and college students, respectively). Moreover, performance on tests of self-derivation through integration relates to academic achievement in both reading and math as indexed by measures that determine grade promotion in kindergarten through Grade 3 (state standardized performance evaluations, classroom grades, and nationally normed assessments). This research suggests that the skills and abilities on which self-derivation of new factual knowledge through integration depends share variance with those that support academic achievement broadly.

Especially in light of evidence of relations between self-derivation through integration task performance and academic achievement, it is important to determine the skills and abilities on which the process depends. Like other productive processes such as analogy, deduction, and induction, self-

derivation through integration may depend on general cognitive abilities such as inhibitory control and working memory (Goswami, 1992, 2011). These abilities tend to be interrelated (e.g., Fukuda, Vogel, Mayr, & Awh, 2010; Nettelbeck & Burns, 2010) and especially during early childhood may also relate to other general cognitive abilities, including abstract reasoning and processing speed (Mungas et al., 2013). Measures of these abilities also relate to academic performance (i.e., high achievers show high performance; e.g., Bhat, 2016; Colom, Escorial, Shih, & Privado, 2007; Floyd, Evans, & McGrew, 2003; Lynn, Meisenberg, Mikk, & Williams, 2007). As such, we would expect positive correlations between measures of these abilities and self-derivation through integration. Consistent with this suggestion, in a study with 4- and 6-year-olds, Varga and Bauer (2014) found that measures of reasoning and verbal comprehension both were correlated with self-derivation performance. Yet only the reasoning task significantly predicted performance in a regression model.

In addition to general cognitive abilities, self-derivation of new factual knowledge through integration may also depend on accumulated semantic knowledge and the long-term retrieval thereof. This possibility is salient because the material over which the cognitive processes are carried out is real-world semantic knowledge. The facts that support self-derivation are previously unknown, yet they relate to concepts about which children have some familiarity (e.g., animals in general, dolphins in particular, how animate objects communicate). As such, long-term retrieval and semantic knowledge could be expected to relate to self-derivation through integration of separate episodes. This logical possibility is supported by the reading comprehension literature that points to a relation between inference making and knowledge (e.g., Cain & Oakhill, 1999). It is further supported by the finding of a correlation between reading comprehension and self-derivation of new factual knowledge through integration in 7- to 11-year-olds (Bauer, Blue, Xu, and Esposito, 2016, Experiment 2). The correlation was observed both when the facts were read by the children themselves and when they were read by an experimenter. Given that reading comprehension is strongly correlated to verbal comprehension and both are considered measures of semantic knowledge and the ability to recruit it, the relation is consistent with suggestions of relations between semantic knowledge and self-derivation through integration.

In summary, self-derivation of new factual knowledge through integration of separate episodes of new learning is a particularly promising area of study due to the ecological validity of the stimuli and paradigm and the relation with academic performance. Yet to date, few studies have examined the relation between self-derivation through integration and other cognitive abilities. The available research is not sufficient to determine what cognitive abilities relate to self-derivation through integration, although the extant research offers support for general cognitive skills such as executive functions and reasoning as well as accumulated knowledge and the ability to use it to reason.

The current research

In the current research, we conducted two experiments to test the skills and abilities on which self-derivation through integration depends. We addressed the question in children across the early school years, when they are heavily engaged in the developmental task of accumulating a knowledge base. In Experiment 1, we assessed self-derivation through integration in 6- and 8-year-olds using a story passage paradigm (Bauer & Larkina, 2017; Bauer & San Souci, 2010). In Experiment 2, we assessed self-derivation through integration in 8- and 10-year-olds using an adaptation of the single-sentence paradigm used by Bauer and colleagues (Bauer et al., 2016; see also Varga & Bauer, 2017, for use of the paradigm with adults). Thus, across the experiments, we examined self-derivation performance in children aged 6–10 years.

In both experiments, children completed a cognitive battery that assessed abstract reasoning, executive functions (including inhibitory control), verbal comprehension, and long-term retrieval. We used the Woodcock–Johnson III Tests of Cognitive Skills (COG; Woodcock, McGrew, & Mather, 2001) to assess reasoning (Concept Formation), verbal comprehension (Verbal Comprehension), and long-term retrieval (Visual Auditory Learning). Given the specific hypotheses regarding inhibition, we added measures of inhibitory function to supplement the Woodcock–Johnson III COG battery. We expected positive associations between self-derivation through integration and reasoning, long-term memory retrieval, and executive functions because self-derivation through integration is a pro-

ductive process and, thus, logically could be expected to follow the same pattern of relation as other such processes, including analogy, deduction, and induction (e.g., Moutier, 2000a; Richland et al., 2006). Based on prior research on self-derivation through memory integration specifically (Varga & Bauer, 2014), we expected relations with abstract reasoning in particular. We also expected positive associations with semantic knowledge because the content over which the processes involved in self-derivation through integration operate is real-world factual knowledge. Previous work has shown a relation between reading comprehension and inference making (e.g., Cain & Oakhill, 1999) and between reading comprehension and self-derivation through integration (Bauer et al., 2016). Given that reading comprehension is a measure of semantic knowledge, it logically could be expected that self-derivation through integration would relate to the depth and breadth of accumulated semantic knowledge and the skills to recruit it to reason on declarative tasks.

Even as we advance these predictions, we note that the cognitive functions under study overlap. Yet given the psychometric structure of the instruments we used to measure them, it is reasonable to anticipate that we could identify relatively unique contributions. Specifically, in factor analyses, the subtests of the Woodcock–Johnson III COG battery (Woodcock et al., 2001) load onto different clusters and, thus, are interpreted to measure separate constructs; Concept Formation loads to the Fluid Reasoning cluster (.81), Verbal Comprehension loads to Comprehension Knowledge (.95), and Visual Auditory Learning loads to Long-Term Retrieval (.71) (Taub & McGrew, 2004). There are no significant residual correlations between the measures. Although the measures of executive functions (including inhibitory function) are not part of the Woodcock–Johnson III COG battery and have not been subjected to factor analyses with the above measures, similar measures from the Delis–Kaplan Executive Function System have been compared with the Woodcock–Johnson III COG battery clusters, with results revealing only moderate correlations, indicating that these measures also represent a separate factor (Floyd et al., 2006). Thus, although we cannot accomplish complete isolation of the different cognitive components under study, we can identify relative unique contributions. The effort is important because understanding the factors that contribute to successful self-derivation through integration can guide the development of interventions and educational methods aimed to facilitate productive knowledge generation.

Experiment 1

Method

Participants

The participants were 20 6-year-olds (9 girls and 11 boys; M_{age} = 6 years 7 months, range = 6;2–6;10 [years;months] and 21 8-year-olds (11 girls and 10 boys; M_{age} = 8 years 6 months, range = 8;2–8;10). Children participated in two sessions approximately 1 week apart, each lasting approximately 1 h. The children were recruited from a volunteer pool consisting of English-speaking families from a large metropolitan area in the southeastern United States who had expressed interest in participating in child development research. Based on parental report, the sample was 24.4% African American, 7.3% biracial or multiracial, and 68.3% Caucasian, with 2% of the sample self-identifying as Hispanic. In terms of education, 5% of reporting parents indicated that the primary caregiver had some college, 2.5% reported having an associate's degree, 40% reported having a college degree, and 52.5% reported having education beyond a bachelor's degree. One family did not report education level. Parents provided informed written consent and children gave verbal assent at the beginning of Session 1. For this and the subsequent experiment, a university institutional review board approved the protocol and procedures. At the end of each session, children received an age-appropriate toy to acknowledge their participation. At the completion of Session 2, parents were given a $10 gift card.

Stimuli

The stimuli were novel "stem" facts from three relational structure groups: "used to make," "located in," and "talks by." There were two different domains within each relational structure, for a total of six unique domains: palm trees and seaweed (used to make), deserts and forests (located

in), and dolphins and hippos (talks by). Within each relational structure, one domain was presented at each testing session, resulting in two sets of similar but unique stimuli. With each domain, each story pair featured two novel stem facts that could be combined to derive a novel integration fact. The stem facts and integration facts from the domains of palm trees, deserts, and dolphins had been used in prior related research (e.g., Bauer & Larkina, 2017); the stem facts and integration facts from the domains of seaweed, forests, and hippos were developed for this research. All facts were accurate and determined by pilot testing to be novel to children in the target age range. Furthermore, pilot testing confirmed that both stem facts were necessary to derive the integration fact (see also Bauer & Larkina, 2017; Bauer & San Souci, 2010).

Each stem fact was presented in the context of a text passage read aloud by an experimenter (see Appendix for an example text passage). The passages were 81–89 words in length distributed over four pages. Each page consisted of a hand-drawn color illustration depicting the main actions of the text; the text was not visible on the page. The passages were similar in structure; in each passage, a character unique to that passage (e.g., ladybug) learned a new fact during the course of an "adventure." Only the stem facts were included in the passages; the integration facts were not presented.

Measures

We assessed a number of cognitive domains.

Reasoning. The Concept Formation subtest of the Woodcock–Johnson III COG (Test 5) was used as a measure of categorical reasoning based on inductive logic (median reliability = .94 from 5 to 99 years) (Woodcock et al., 2001). Prior research has revealed relations between abstract reasoning and self-derivation through integration (Varga & Bauer, 2014). Participants saw a row of shape and color patterns and were required to derive the rule that governs each sequence. Participants received 1 point for each correctly answered item; points were summed for a total score. The total score was then used to determine the standardized score using the Woodcock–Johnson III Compuscore and Profiles program. The resulting score was based on children's performance compared with a comparison group of same-aged peers, with a mean of 100 and a standard deviation of 15.

Executive functions. Executive functions were measured with three tasks: the Trail Making Task, Bivalent Shape Task, and Go/No-Go task. All three were administered on a Dell Optiplex 755 computer attached to a Dell S2240Tb 21.5-inch color touchscreen monitor using the Psychology Experiment Building Language (Mueller, 2010). These tasks were included due to relations between executive functions and other productive processes (Goswami, 1992, 2011).

We administered a version of the Trail Making Task, a measure of cognitive flexibility, developed specifically for children (e.g., Bowie & Harvey, 2006; Delis, Kaplan, & Krames, 2001; Mueller, 2010; Reitan, 1971). Children completed three trails by putting items in order. The first trail required numeration from 1 to 16 (1–2–3). In the second trail, the sequence was alphabetical from A to K (A-B-C). In the last trail, the sequence alternated between a number and a letter (1-A-2-B-3-C). Response time for the last trail is the measure of cognitive flexibility. The Bivalent Shape Task, a measure of inhibitory control, requires children to match the shape between objects and ignore the highly salient color (Esposito, Baker-Ward, & Mueller, 2013). The dependent variable was mean reaction time on correct incongruent trials (color and shape mismatch). The Go/No-Go paradigm (Archibald & Kerns, 1999), a second measure of inhibitory control, was adapted for touchscreen use (Mueller, 2010). Participants were asked to tap the screen in response to the target stimulus. The target stimulus appeared with 80% frequency and nontarget with 20% frequency. The dependent variable taken from this measure was total errors.

Long-term retrieval. The Visual Auditory Learning subtest of the Woodcock–Johnson III COG (Test 2 and Test 10) served as a measure of long-term memory for associated stimuli (median reliability = .86 from 5 to 99 years) (Woodcock et al., 2001). This measure was included to ascertain whether self-derivation through integration is related to the storing and reactivation of information in long-term memory. In this task, participants are shown a series of rebuses (pictographic symbols of words) and are later asked to recall the visual–auditory associations from long-term memory. Specifically,

once the rebuses are learned, participants are presented with several rebuses forming a sentence and are asked to speak the associated words aloud (Test 2). One week later, participants were presented with new series of rebuses to "read" without being reminded of the word associations (Test 10). In both Test 2 and Test 10, participants received 1 point for each incorrectly answered item, defined as a failure to identify the correct word or to do so within 5 s of viewing a rebus. Total errors were entered into the Woodcock–Johnson III Compuscore and Profiles program, which returned standard scores for each measure based on children's performance compared with same-aged peers (again, a mean of 100 and a standard deviation of 15).

Verbal comprehension. The Woodcock–Johnson III Test 1 is a norm-referenced verbal comprehension measure (median reliability = .90 from 5 to 99 years) (Woodcock et al., 2001). The test includes four subtests—Picture Vocabulary, Synonyms, Antonyms, and Analogies—which assess comprehension of individual words as well as the relation among words. The test represents context-free knowledge and sematic memory (Schrank, 2010). This measure was included to ascertain the relation between self-derivation through integration and acquired semantic knowledge and is based on evidence of relations between self-derivation and academic performance and reading comprehension (Bauer et al., 2016; Esposito & Bauer, 2017). Children received 1 point for each correctly answered item in each subtest. The subtest scores were then entered into the Woodcock–Johnson III Compuscore and Profiles program, which returned a composite standard score based on children's performance compared with same-aged peers (again, a mean of 100 and a standard deviation of 15).

Procedure

The procedure was modeled after related research (Bauer & Larkina, 2017; Bauer & San Souci, 2010). Children were tested individually in a laboratory room by one of two female experimenters; children were tested by the same experimenter at each of the two sessions. Each session was approximately 1 h long. Sessions were video-recorded, and the recordings were regularly reviewed to ensure protocol fidelity.

Session 1. At the first session, children were exposed to three story passage pairs containing related facts, one story passage pair from each of the relational structure groups (i.e., used to make, located in, and talks by). Each story passage was read twice while children were encouraged to listen and look at the pictures. The first passage from each pair was read, followed by verbal comprehension Subtests A and B. The second passage from each story pair was then read. Following the second passage, children participated in the computerized measures of executive functions in random order. Children were then tested for self-derivation through integration performance in an open-ended format. Following open-ended testing, for trials on which children did not provide a correct answer in open-ended testing, they were presented with a forced-choice question with three alternatives, one of which was correct. Forced-choice testing was included to ensure that there would be sufficient variability if open-ended testing resulted in a restricted range. We concluded the session with the long-term retrieval task.

Session 2. Children returned to the laboratory approximately 1 week after their initial visit (M = 7.2 days). The session proceeded in much the same way as Session 1. Previously unseen passage pairs from each relational structure were presented. Story passages within each pair were separated by the verbal comprehension Subtests C and D. Computerized executive function tasks again separated the passages from self-derivation through integration performance. Children were first asked self-derivation questions in an open-ended format followed by a forced-choice format (again, only when children did not produce a correct answer in open-ended testing). After the tests, children completed the reasoning measure and the delayed long-term retrieval measure. For the self-derivation through integration task, across participants, story passage pairs were presented approximately equally often in Session 1 and Session 2. In addition, within a story passage pair, stories were presented first and second approximately equally often across participants.

Scoring of self-derivation through integration

The experimenters recorded children's answers to both the open-ended and forced-choice integration questions. Participants were awarded 1 point for each correctly provided answer in the open-ended and forced-choice formats. Thus, there were two measures of performance: open-ended and total score (calculated by adding open-ended and forced-choice scores). Scoring was checked by each experimenter to ensure fidelity.

Results

The results are reported in two sections. First, we report the self-derivation through integration performance and the relation between sessions. Second, we report the relations between measures of cognitive abilities and self-derivation through integration performance. Analyses were conducted using SPSS Version 24. All statistical tests were two-tailed.

Self-derivation performance

Self-derivation through integration performance for both Sessions 1 and 2 is represented in Fig. 1 for both 6- and 8-year-olds. Overall, the 6-year-olds showed evidence of self-derivation through integration at both sessions without ceiling or floor effects. The 8-year-olds showed variability in open-ended performance at both Sessions 1 and 2 but reached ceiling in the total score at both sessions. Performance from Session 1 to Session 2 did not differ significantly for either 6-year-olds (open-ended: $t = -1.19$, $p = .25$; total score: $t = -1.19$, $p = .65$) or 8-year-olds (open-ended: $t = -0.96$, $p = .35$; total score: $t = 1.45$, $p = .16$). Across the sample, Session 1 and Session 2 were significantly correlated ($r = .45$, $p = .004$). Given that there were no differences in performance and that performance across the two sessions was correlated, Session 1 performance and Session 2 performance were summed, effectively increasing the range from 0 to 3 to 0 to 6. In addition, given that there was sufficient variance in open-ended testing ($M = 3.51$, $SD = 1.86$, 59% correct, range = 6) but not in total score ($M = 5.28$, $SD = 1.34$, 88% correct, range = 5), only open-ended performance was examined further.

Relations with cognitive abilities

Performance on the measures of cognitive abilities is reported in Table 1, Panel A. In Table 2, Panel A, we report the correlations. Reasoning and verbal comprehension were significantly correlated to self-derivation through integration performance. No other cognitive abilities were correlated to self-derivation performance.

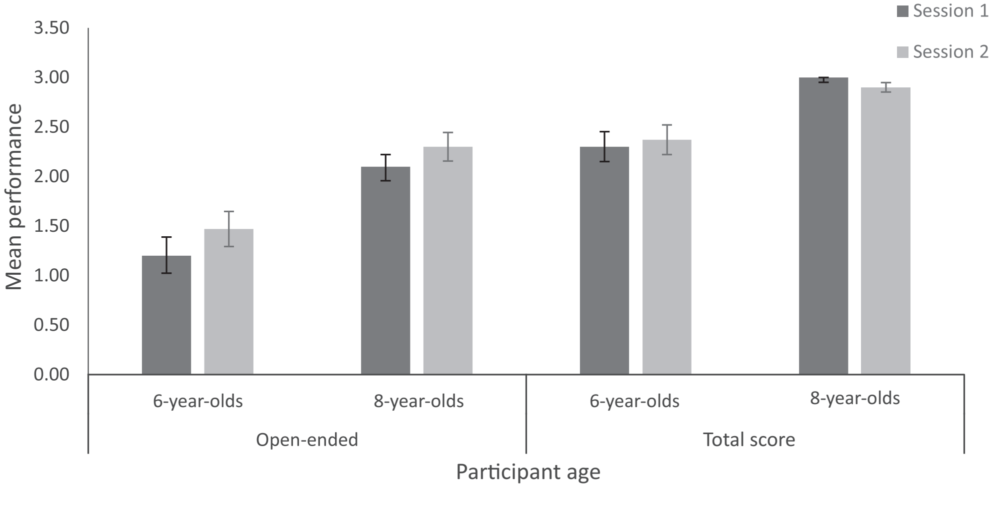


Fig. 1. Mean self-derivation performance across test sessions: Experiment 1.

Table 1
Performance on cognitive measures.

	Panel A: Experiment 1	Panel B: Experiment 2
Verbal comprehension, standardized score	115.49 (14.35)	115.58 (13.87)
Long-term memory, standardized score	112.93 (15.91)	104.15 (11.90)
Long-term memory delay, standardized score	118.74 (13.53)	103.87 (10.58)
Reasoning, standardized score	107.29 (13.53)	108.47 (11.87)
TMT, total completion time	48.48 (29.31)	35.60 (14.08)
BST, mean reaction time	1.14 (0.21)	1.05 (0.19)
GNG, mean errors	4.37 (2.60)	3.68 (1.61)

Note. Standard deviations are in parentheses. TMT, Trail Making Task; BST, Bivalent Shape Task; GNG, Go/No-Go task.

We next conducted a regression analysis to ascertain whether reasoning and verbal comprehension, the measures that were significantly correlated with self-derivation performance, explained unique variance (see Table 3). Age of participants was also included in the model. The model was significant and explained 51% of the variance in self-derivation through integration performance. Only age and verbal comprehension were significant predictors of performance; reasoning did not explain a significant portion of the variance.

Discussion

Both 6- and 8-year-olds showed evidence of self-derivation of new knowledge through integration of separate yet related episodes of new learning. Performance at Session 1 and performance at Session 2 were not significantly different, and performance at the two sessions was significantly correlated.

Of the measures of general cognitive abilities, only reasoning and verbal comprehension were correlated to self-derivation through integration performance. Yet in the regression model, reasoning did not predict unique variance. Instead, verbal comprehension—a measure of the depth and breadth of accumulated semantic knowledge—and participant age explained significant and unique variance.

The method used in the current experiment afforded a valid test of the question of the specificity of relations between measures of cognitive and semantic development and self-derivation of new knowledge through integration. Yet one limitation of the approach was that we measured children's performance on only 6 trials, yielding a somewhat restricted range of scores. In open-ended testing, the performance range was 6 and in total score the range was 5. Accordingly, in Experiment 2 we adopted a new method of testing that more than quadrupled the number of trials in each session. We also increased the age range to include 10-year-olds. We did not include 6-year-olds, however, in anticipation that they would not be able to meet the requirement that they read the stimulus facts themselves (see below).

Experiment 2

Method

Participants

The participants were 21 8-year-olds (8 girls and 13 boys; M_{age} = 8 years 5 months, range = 8;2–8;11) and 19 10-year-olds (9 girls and 10 boys; M_{age} = 10 years 5 months; range = 10;1–10;10). Children participated in two sessions, each 1 h long, approximately 1 week apart. Children were recruited from the same source and represent the same population as in Experiment 1. None of the children had participated in Experiment 1. Based on parental report, the sample was 25% African American, 18% biracial or multiracial, 3% Asian, and 50% Caucasian, with 3% of the sample self-identifying as Hispanic. Two parents did not respond to the request for racial self-classification. In terms of education, 15% percent of parents reported the primary caregiver having a high school degree, 10% reported having an associated degree, 22.5% reported having a college degree, and 52.5% reported having education beyond a bachelor's degree. As in Experiment 1, parents provided

Table 2
Correlations between self-derivation through integration and cognitive measures.

	1. Self-derivation	2. Verbal comprehension	3. Long-term memory	4. Long-term memory delay	5. Reasoning	6. TMT	7. BST	8. GNG
Panel A: Experiment 1								
1. Open-ended self-derivation								
2. Verbal comprehension, standardized score	$.45^{**}$							
3. Long-term memory, standardized score	.28	$.33^{*}$						
4. Long-term memory delay, standardized score	.17	.30	$.77^{**}$					
5. Reasoning, standardized score	$.33^{*}$	$.59^{**}$	.25	.18				
6. TMT, total completion time	.20	−.21	.15	.05	−.002			
7. BST, mean reaction time	−.05	−.007	−.02	.09	−.20	−.002		
8. GNG, mean errors	−.006	.01	.11	.001	.02	.30	−.007	
Panel B: Experiment 2								
1. Open-ended self-derivation								
2. Verbal comprehension, standardized score	$.42^{**}$							
3. Long-term memory, standardized score	.17	.30						
4. Long-term memory delay, standardized score	.05	.15	$.59^{**}$					
5. Reasoning, standardized score	$.39^{*}$	$.44^{**}$	$.45^{**}$	$.58^{**}$				
6. TMT, total completion time	−.22	−.08	.14	$.41^{*}$	.007			
7. BST, mean reaction time	−.06	.29	.07	.13	.18	.28		
8. GNG, mean errors	−.11	−.11	.002	.16	.005	.23	−.16	

Note. TMT, Trail Making Task; BST, Bivalent Shape Task; GNG, Go/No-Go task.
* $p < .05$.
** $p < .01$.

Table 3
Regression model predicting self-derivation performance: Experiment 1.

		ΔR^2	β	*VIF*
	Predictor			
	Age		.54**	1.06
	Reasoning		−.07	1.08
	Verbal comprehension		.49*	1.12
Total R^2		.51**		
n		38		

Note. VIF, variance inflation factor.

informed written consent and children gave verbal assent at the beginning of Session 1. At the end of each hour-long session, children received an age-appropriate toy to acknowledge their participation. At the completion of Session 2, parents were given a $10 gift card.

Stimuli

The stimuli were 60 novel "stem" facts in 30 related pairs that could be integrated to create 30 novel "integration" facts. Preliminary laboratory testing revealed that both stem facts were necessary for production of the integration facts and that the facts were novel to children in the target age range (see also Bauer et al., 2016). Unlike in Experiment 1, the facts were presented as individual sentences rather than in the context of an illustrated story passage (see Bauer et al., 2016, for use of a similar approach; see Appendix for example stimuli). The sentences were presented in written form via PowerPoint. To encourage engagement, after approximately every fourth fact, children were asked to categorize the previously presented fact. For example, after seeing and hearing "Titan is Saturn's largest moon," children were asked, "Was this fact about planets or animals?" Answers to the engagement questions were not analyzed. Two lists of 15 pairs were created from the 30 pairs. The lists were pilot tested and showed equivalent performance.

Measures

The measures were the same as in Experiment 1.

Procedure

The procedure resembled Experiment 1, with the exception that integration facts were presented via computer in the single-sentence paradigm. Children were tested individually in a laboratory room by one of two female experimenters; children were tested by the same experimenter at each of the two sessions. Both of the hour-long sessions were video-recorded, and the recordings were regularly reviewed to ensure protocol fidelity.

Session 1. At the first session, children were exposed to 15 sentence pairs containing related facts, either List A or List B. Children read the facts aloud, thereby permitting us to monitor that the facts were read in their entirety. As in Experiment 1, following the presentation of the first stem facts from each stem fact pair, children completed the verbal comprehension Subtests A and B. After the second stem facts from each pair were presented, children completed the executive function tasks. Children were then tested for self-derivation through integration performance in an open-ended format. Again, to ensure sufficient variability for analyses if open-ended testing produced a restricted range, children were next tested in a forced-choice format. The forced-choice questions were presented on the computer; thus, all children answered all forced-choice questions rather than only questions they were not successful in answering in the open-ended format. We concluded the session with the long-term retrieval task.

Session 2. Children returned to the laboratory approximately 1 week after their initial visit (M = 7.3 days). The session modeled Session 2 of Experiment 1. The previously unseen fact list (A or B) was

presented. As in Session 1, the first member of each stem fact pair was presented. Children then completed verbal comprehension Subtests C and D. Following that, children were presented with the second member of each stem fact pair. Computerized executive function tasks separated the facts from the test for self-derivation through integration. Children were tested first in an open-ended format followed by a forced-choice format, using the same procedure as in Session 1. Across participants, Lists A and B were counterbalanced across Sessions 1 and 2. In addition, within Lists A and B, the fact order presentation was counterbalanced. After test, children completed the reasoning measure and the delayed long-term retrieval measure.

Scoring

The experimenters recorded answers provided during the session. Participants were awarded 1 point for each correct answer for open-ended and forced-choice testing. Unlike Experiment 1, all children were asked all forced-choice questions, eliminating the need for a "total score" measure. Thus, there were two measures of performance: open-ended and forced choice. Scoring was checked by each experimenter to ensure fidelity.

Results

The results are again reported in two sections: the self-derivation through integration performance and the relation between sessions and then the relation among measures of general cognitive abilities, semantic knowledge, and self-derivation through integration performance. Analyses were conducted using SPSS Version 24. All statistical tests are two-tailed.

Self-derivation performance

Self-derivation through integration performance for both Session 1 and Session 2 is represented in Fig. 2 for both 8- and 10-year-olds. Overall, both age groups showed evidence of self-derivation through integration across sessions with sufficient variability to indicate no floor or ceiling effects in either open-ended or forced-choice testing. Performance in 8-year-olds did not differ significantly across Session 1 and Session 2 testing (open-ended: $t = 1.87$, $p = .08$; forced choice: $t = 0.81$, $p = .43$). Similarly, 10-year-olds' performance did not differ significantly across the sessions (open-ended: $t = 1.36$, $p = .19$; forced choice: $t = 1.07$, $p = .30$). Across the sample, Session 1 performance was significantly correlated to Session 2 performance, $r = .67$, $p < .001$ (open-ended: $M = 14.65$, $SD = 5.96$, 49% correct, range = 21), and $r = .44$, $p = .005$ (forced choice: $M = 23.58$, $SD = 4.45$, 79% correct,

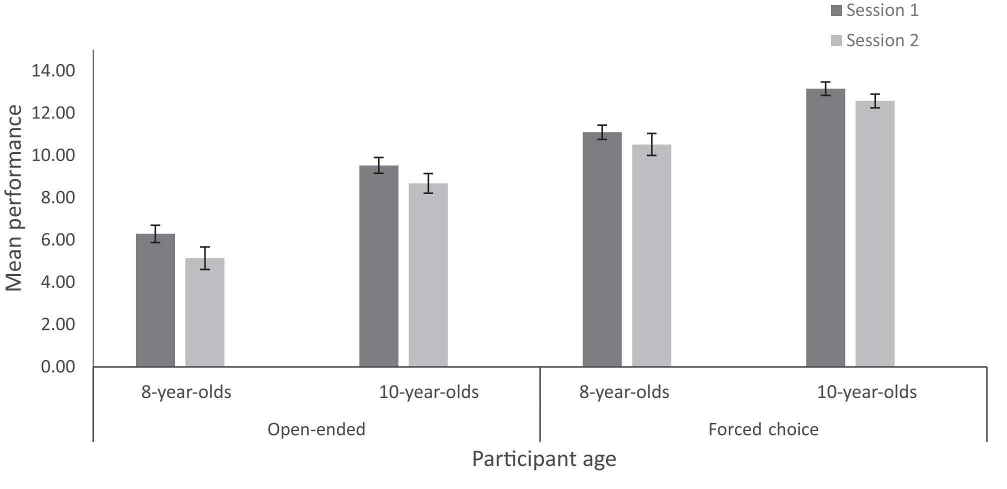


Fig. 2. Mean self-derivation performance across test sessions: Experiment 2.

Table 4
Regression model predicting self-derivation performance: Experiment 2.

		ΔR^2	β	*VIF*
	Predictor			
	Age		.45*	1.13
	Reasoning		.12	1.27
	Verbal comprehension		.39*	1.41
Total R^2		.45**		
n		40		

Note. VIF, variance inflation factor.

range = 20). Given that open-ended performance provided sufficient variability, we proceeded with analyses using this measure only.

Relations with cognitive abilities

Performance on the cognitive ability measures are reported in Table 1, Panel B. In Table 2, Panel B, we report the correlations between self-derivation through integration and the measures of cognitive abilities. Following the same pattern as Experiment 1, reasoning and verbal comprehension were significantly correlated to self-derivation through integration performance. No other cognitive abilities were correlated to self-derivation performance.

We used a regression analysis to ascertain whether reasoning and verbal comprehension both explained unique variance in self-derivation performance (see Table 4). Age was also included. The model accounted for 45% of the variance in self-derivation through integration performance. Both age and verbal comprehension were significant. As in Experiment 1, reasoning was not a significant predictor of performance.

Discussion

In Experiment 2, we introduced a new paradigm for measuring self-derivation through integration of separate but related episodes of learning. Both 8- and 10-year-olds were successful in this paradigm and showed sufficient variability in both open-ended format and forced-choice testing. Performance was correlated across sessions.

As in Experiment 1, both reasoning and verbal comprehension were correlated to self-derivation through integration performance. Also as in Experiment 1, only verbal comprehension explained significant unique variance in the regression analyses in addition to age. The measures of long-term memory, long-term memory with a week-long delay, and executive functions did not correlate with self-derivation performance. Thus, the relation between self-derivation through integration and general reasoning was not as strong as the relation with acquired semantic knowledge.

General discussion

In the current research, we examined the relation of self-derivation of new factual knowledge through integration of separate yet related episodes with tests of component cognitive abilities. The aim was to determine with which component cognitive abilities self-derivation shares variance. Associations with reasoning, long-term memory retrieval, and executive functions were expected because self-derivation through integration is a productive process and, thus, logically could be expected to follow the same pattern of relation as other such processes, including analogy, deduction, and induction (e.g., Moutier, 2000a; Richland et al., 2006). This expectation was bolstered by findings that self-derivation through integration predicts academic achievement in both reading and math (Esposito & Bauer, 2017), a pattern observed for other general cognitive abilities across the school years (e.g., Bhat, 2016; Floyd et al., 2003; Spinath, Spinath, Harlaar, & Plomin, 2006) and in adults (e.g., Rohde &

Thompson, 2007; Ruffling, Wach, Spinath, Brünken, & Karbach, 2015). At the same time, we also expected associations with semantic knowledge because the material over which the processes involved in self-derivation through integration operate is real-world factual knowledge. Previous work has shown a relation between reading comprehension and inference making (e.g., Cain & Oakhill, 1999) and between reading comprehension and self-derivation through integration (Bauer et al., 2016). Given that reading comprehension is a measure of semantic knowledge, it logically could be expected that self-derivation through integration would relate to the depth and breadth of accumulated semantic knowledge and the skills to recruit it to reason on declarative tasks.

Despite differences between the studies in age of participants, stimuli, and protocols, both experiments yielded the same pattern of findings. In Experiment 1, children aged 6 and 8 years listened to three story pairs in each of two sessions. The combined performance across the sessions was correlated to both a measure of reasoning and a measure of verbal comprehension. However, in a subsequent regression model, only child age and verbal comprehension predicted unique variance in self-derivation performance. In Experiment 2, 8- and 10-year-old children read 15 individual fact pairs in each of two sessions. The combined performance was again correlated to both a measure of reasoning and a measure of verbal comprehension. Again, only age and verbal comprehension predicted unique variance in self-derivation performance. The results of both studies, thus, indicate that self-derivation of new factual knowledge through integration shares variance with abstract reasoning but that the more significant determinant of performance is the ability to acquire and recruit semantic knowledge in the service of a specific task demand. Although certainly these cognitive functions overlap, given the psychometric structure of the instruments we used to measure them, it is reasonable to conclude unique contributions. Specifically, we used measures from the Woodcock–Johnson III COG (Woodcock et al., 2001). In factor analyses, the measures load onto different clusters and, thus, are interpreted to measure separate constructs; Verbal Comprehension loads to Comprehension Knowledge (.95), and Concept Formation loads to the Fluid Reasoning cluster (.81) (Taub & McGrew, 2004). There are no significant residual correlations between the measures. As such, the results of the current research can be taken to indicate that the ability to accumulate and recruit semantic knowledge is uniquely predictive of self-derivation through integration.

The findings of the current research are important methodologically, and they have implications for both theory and practice. From a methodological standpoint, the current research provides converging evidence across numerous task differences. That is, although different tasks were used in Experiment 1 and Experiment 2, the pattern of results was the same. First, in both experiments, open-ended task performance featured sufficient variability for valid tests of the study hypotheses. Second, in both experiments, performance was correlated across the two sessions. This indicates test–retest reliability in both the story passage paradigm and the single-sentence paradigm, at least for the ages tested in these experiments. Third, performance did not significantly differ from Session 1 to Session 2 in either experiment. This indicates that the paradigms can be implemented with their respective age groups in sessions that are at least 1 week apart without concern regarding substantial practice effects. Fourth, the new single-sentence paradigm used in Experiment 2 quadrupled the number of trials than can be tested using the story passage paradigm and showed variance in both open-ended and forced-choice testing with both 8- and 10-year-olds. This new paradigm, thus, holds promise for use with children in this age range and older, permitting a methodological means to bridge the gap between the child literature (e.g., Bauer & Larkina, 2017) and adult literature (Bauer & Jackson, 2015; Varga & Bauer, 2017) on self-derivation of new factual knowledge through memory integration.

The finding that the pattern of results was the same across Experiment 1 and Experiment 2, in spite of the fact that children's ages spanned 6–10 years, also has important theoretical implications regarding the status of abstract reasoning ability relative to other cognitive abilities. Studies associated with the validation of the National Institutes of Health Toolbox Cognition Battery revealed that across the age range of 3–15 years, the measure of verbal comprehension represented a separate dimension of cognition (Mungas et al., 2013). In contrast, measures of attention, episodic memory, executive functions, speed of processing, and working memory exhibited different structures in younger and older children (measures of reasoning were not included in the Cognition Battery). Whereas during later childhood (8–15 years) the abilities differentiated from one another, during early childhood (3–6 years) the measures all loaded on the same dimension. This finding has been taken to suggest that dif-

ferentiation of cognitive abilities is a developmental achievement (Mungas et al., 2013). The fact that in the current research only the measure of abstract reasoning—and not the measures of episodic memory and executive functions—was correlated with self-derivation through integration performance implies that this conclusion should be qualified. It seems that even at 6 years of age reasoning ability is at least partially differentiated from other general cognitive abilities.

The findings of the current research also have important implications for efforts to promote reasoning skills and abilities and productive processes more broadly. The developmental and adult literatures are replete with studies in which cognitive skills and abilities are trained, followed by a test for transfer to academic or other real-world contexts. Such interventions typically show improvements in performance on the materials used for training, some improvement on materials similar to those used for training (near transfer), but disappointing levels of transfer to the academic and everyday cognitive domains the training was intended to improve (far transfer; e.g., Dunning, Holmes, & Gathercole, 2013; Titz & Karbach, 2014; see also Melby-Lervåg, Redick, & Hulme, 2016; see Simons et al., 2016, for meta-analyses; see Serpell & Esposito, 2016, for review and discussion). For example, Mackey, Park, Robinson, and Gabrieli (2017) provided cognitive-based training to fifth-grade students. The stimuli were playing cards with arbitrary images. Each of four games targeted a different general cognitive ability: speed of processing, working memory, reasoning, and task switching. The interventions were implemented 20 min per day, 4 days per week, for 9 weeks. At the end of the school year during which the training took place, students in the training condition had improved performance in areas of training compared with a control group that received typing and geography training. However, there were no between-group differences in academic outcomes. The findings of the current research suggest one possible explanation for poor transfer observed in this exemplar and related studies, namely that whereas training studies frequently target general cognitive skills and abilities, the productive processes they are intended to improve are performed over real-world factual content. As revealed in the current research, under such circumstances the processes seem to draw more heavily from semantic knowledge than from general cognitive abilities. This suggests that interventions that more closely approximate self-derivation of new factual knowledge through integration of separate yet related episodes of new learning may be more effective than those that employ abstract problems and arbitrary associations.

The results of the current research are consistent with those of Bauer et al. (2016), which indicated relations between self-derivation through integration and reading comprehension in 7- to 11-year-old children. At the same time, they are not fully consistent with the results of Varga and Bauer (2014) with children aged 4 and 6 years. As in the current research, Varga and Bauer found that both reasoning and verbal comprehension were related to self-derivation through integration. However, only reasoning predicted significant and unique variance in self-derivation performance. We speculate that the different relative weighting of reasoning and verbal comprehension among younger children (average age 5 years) and older children (average age 8 years) reflects a developmental progression. It is reasonable to assume that the younger sample in Varga and Bauer (2014) had more limited accumulated semantic knowledge as well as less skill recruiting it in the service of a task demand. Indeed, semantic knowledge shows rapid development related to formalized schooling, a period including ages 6–10 years (e.g., Cahan & Cohen, 1989). With limited semantic knowledge to bring to bear, young children may be more dependent on general cognitive abilities (Goswami, 2011). We may further speculate that less well-developed semantic knowledge coupled with less well-developed reasoning abilities in 4-year-olds relative to 6-year-olds may help to explain the substantial difference in open-ended self-derivation performance between these age groups (i.e., 13% and 50–65% for 4- and 6-year-olds, respectively; Bauer & Larkina, 2017; Bauer & San Souci, 2010). It also may help to explain 4-year-olds' substantially better performance in forced-choice testing relative to open-ended testing (62% and 13%, respectively; Bauer & San Souci, 2010; see also Bauer & Larkina, 2017). In forced-choice testing, children are not responsible for retrieving and marshalling the necessary information themselves (as they are in open-ended testing) but instead can select the correct response from among alternatives provided to them.

The experiments described here are not without limitations that provide direction for future work. The story passage paradigm used in Experiment 1 had a small range (0–6). This was largely remedied in Experiment 2, with a range that increased to 0–30. In addition, the experiments both took place in a

laboratory environment. Future research should move into more ecologically valid learning environments such as classrooms and museums (see Esposito & Bauer, 2017, for steps in this direction). The experiments also took place over a 1-week period, eliminating the possibility of exploring cross-lagged relations. Future research should explore the cross-lagged relations with a range of cognitive abilities over a longer time frame. Such work would provide information regarding the directionality of the relation between self-derivation of new knowledge through integration performance and semantic knowledge and how this relation unfolds over development.

Conclusions

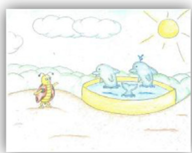
The current research suggests that self-derivation through integration of separate episodes is predicted by accumulated knowledge and the ability to use the knowledge in the service of a specific task demand. Although self-derivation shared variance with reasoning skills, reasoning did not predict unique variance in individual performance. The results provide a possible explanation for lack of transfer in cognitive training studies focused on increasing general cognitive abilities through the use of arbitrary stimuli. Perhaps one answer to improving means to build a semantic knowledge base is to design tasks that require the recruitment and use of that knowledge.

Acknowledgments

The work was supported by grants from the National Institute of Child Health and Human Development (HD067359), National Science Foundation (BCS1528091), and Institute of Education Sciences (R305A150492) to Patricia J. Bauer and by the Emory College of Arts and Sciences. We thank members of the Bauer Memory at Emory laboratory group for their assistance with data collection and help with various aspects of the research. We also thank the children and parents who took part in the research.

Appendix

Story passage example

Passage	Page 1	Page 2	Page 3	Page 4
Ladybug goes to the Zoo				
	One day, a ladybug went to the zoo so that she could make some new friends.	At the zoo, she met some friendly dolphins playing in the water. “Friendly dolphins,” she asked, “may I be part of your group?	The dolphins said, “We’d love to have you join our pod. But you’ll have to live in the water with us.	The ladybug shook her head sadly and then she left to go home. But now she knew that a group of dolphins was called a pod.

Single-sentence game example

Stem Fact 1	Intervening facts and activities	Stem Fact 2	Intervening facts and activities	Integration Question
Titan is Saturn's largest moon.		Saturn's largest moon is the only moon with clouds.		What is the name of the only moon with clouds? (*Titan*)

References

Akshoomoff, N., Beaumont, J. L., Bauer, P. J., Dikmen, S. S., Gershon, R. C., Mungas, D., ... Heaton, R. K. (2013). NIH Toolbox Cognition Battery (CB): Composite scores of crystallized, fluid, and overall cognition. *Monographs of the Society for Research in Child Development, 78*, 119–132.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology, 5*, 115–129.

Bauer, P. J. (2009). Neurodevelopmental changes in infancy and beyond: Implications for learning and memory. In O. A. Barbarin & B. H. Wasik (Eds.), *Handbook of child development and early education: Research to practice* (pp. 78–102). New York: Guilford.

Bauer, P. J. (2012). Facilitating learning and memory in infants and young children: Mechanisms and methods. In S. L. Odom, E. P. Pungello, & N. Gardner-Neblett (Eds.), *Infants, toddlers, and families in poverty: Research implications for early child care* (pp. 49–74). New York: Guilford.

Bauer, P. J., Blue, S. N., Xu, A., & Esposito, A. G. (2016). Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources. *Developmental Psychology, 52*, 1024–1037.

Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 271–282.

Bauer, P. J., & Larkina, M. (2017). Realizing relevance: The influence of domain-specific information on generation of new knowledge through integration in 4- to 8-year-old children. *Child Development, 88*, 247–262.

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology, 107*, 452–465.

Bauer, P. J., & Varga, N. L. (2017). Similarity and deviation in event segmentation and memory integration: Commentary on Richmond, Gold, & Zacks. *Journal of Applied Research in Memory and Cognition, 6*, 124–128.

Bhat, M. A. (2016). The predictive power of reasoning ability on academic achievement. *International Journal of Learning, Teaching and Educational Research, 15*(1), 79–88.

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols, 1*, 2277–2281.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1982). *Learning, remembering, and understanding. Technical Report No. 244*. Center for the Study of Reading, University of Illinois at Urbana-Champaign.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development, 60*, 1239–1249.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*, 489–503.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*, 31–42.

Colom, R., Escorial, S., Shih, P. C., & Privado, J. (2007). Fluid intelligence, memory span, and temperament difficulties predict academic performance of young adolescents. *Personality and Individual Differences, 42*, 1503–1514.

Delis, D. C., Kaplan, E., & Krames, J. H. (2001). Trail Making Task. In *Delis-Kaplan executive function system: Examiner manual* (pp. 39–54). San Antonio, TX: Psychological Corporation.

Dunning, D. L., Holmes, J., & Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working memory? A randomized controlled trial. *Developmental Science, 16*, 915–925.

Esposito, A. G., Baker-Ward, L., & Mueller, S. T. (2013). Interference suppression vs. response inhibition: An explanation for the absence of a bilingual advantage in preschoolers' Stroop task performance. *Cognitive Development, 28*, 354–363.

Esposito, A. G., & Bauer, P. J. (2017). Going beyond the lesson: Self-generating new factual knowledge in the classroom. *Journal of Experimental Child Psychology, 153*, 110–125.

Fisher, A. V., Godwin, K. E., & Matlen, B. J. (2015). Development of inductive generalization with familiar categories. *Psychonomic Bulletin & Review, 22*, 1149–1173.

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell–Horn–Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools, 40*, 155–171.

Floyd, R. G., McCormack, A. C., Ingram, E. L., Davis, A. E., Bergeron, R., & Hamilton, G. (2006). Relations between the Woodcock-Johnson III clinical clusters and measures of executive functions from the Delis-Kaplan Executive Function System. *Journal of Psychoeducational Assessment, 24*, 303–317.

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity not quality: The relationship between fluid intelligence and working memory capacity. *Psychological Bulletin & Review, 17*, 673–679.

Goswami, U. (1992). *Analogical reasoning in children*. New York: Psychology Press.

Goswami, U. (2011). Inductive and deductive reasoning. In U. Goswami (Ed.), *Childhood cognitive development* (pp. 399–419). Oxford, UK: Wiley-Blackwell.

Lynn, R., Meisenberg, G., Mikk, J., & Williams, A. (2007). National IQs predict differences in scholastic achievement in 67 countries. *Journal of Biosocial Science, 39*, 861–874.

Mackey, A. P., Park, A. T., Robinson, S. T., & Gabrieli, J. D. (2017). A pilot study of classroom-based cognitive skill instruction: Effects on cognition and academic performance. *Mind, Brain, and Education, 11*(2), 85–95.

Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: Inductive generalization in infancy. *Cognition, 59*, 307–335.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer" evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*, 512–534.

Moutier, S. (2000b). Deductive reasoning and experimental matching–bias inhibition training in school children. *Cahiers de Psychologie Cognitive, 19*, 429–452.

Moutier, S. (2000a). Deductive competence and executive efficiency in school children. *Current Psychology Letters, 3*, 87–100.

Moutier, S., Plagne-Cayeux, S., Melot, A., & Houdé, O. (2006). Syllogistic reasoning and belief-bias inhibition in school children: Evidence from a negative priming paradigm. *Developmental Science, 9*, 166–172.

Mueller, S. T. (2010). The psychology experiment building language (Version 0.11) [computer software].

Mungas, D., Widaman, K., Zelazo, P. D., Tulsky, D., Heaton, R. K., Slotkin, J., ... Gershon, R. C. (2013). NIH Toolbox Cognition Battery (CB): Factor structure for 3- to 15-year-olds. *Monographs of the Society for Research in Child Development, 78*, 103–118.

Nettelbeck, T., & Burns, N. R. (2010). Processing speed, working memory and reasoning ability from childhood to old age. *Personality and Individual Differences, 48*, 379–384.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*, 443–468.

Perret, P. (2015). Children's inductive reasoning: Developmental and educational perspectives. *Journal of Cognitive Education and Psychology, 14*, 389–408.

Reitan, R. M. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33*, 575–581.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249–273.

Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42*, 546–567.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83–92.

Ruffling, S., Wach, F.-S., Spinath, F. M., Brünken, R., & Karbach, J. (2015). Learning strategies and general cognitive ability as predictors of gender-specific academic achievement. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01238.

Schrank, F. A. (2010). Woodcock-Johnson III Tests of Cognitive Abilities. In A. S. Davis & R. C. D'Amato (Eds.), *Handbook of pediatric neuropsychology* (pp. 415–434). New York: Springer.

Serpell, Z. N., & Esposito, A. G. (2016). Development of executive functions: Implications for educational policy and practice. *Policy Insights from the Behavioral and Brain Sciences, 3*, 203–210.

Siegler, R. S. (1989). Mechanisms of cognitive development. *Annual Review of Psychology, 40*, 353–379.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17*(3), 103–186.

Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence, 34*, 363–374.

Taub, G. E., & McGrew, K. S. (2004). A confirmatory factor analysis of Cattell–Horn–Carroll theory and cross-age invariance of the Woodcock-Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly, 19*(1), 72–87.

Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research, 78*, 852–868.

Varga, N. L., & Bauer, P. J. (2013). Effects of delays on 6-year-old children's self-generation and retention of knowledge through integration. *Journal of Experimental Child Psychology, 115*, 326–341.

Varga, N. L., & Bauer, P. J. (2014). Conceptual knowledge extension: An examination of its development and the underlying cognitive processes involved. In R. Chen (Ed.), *Cognitive development: Theories, stages and processes and challenges* (pp. 1–16). Huntington, NY: Nova Science.

Varga, N. L., & Bauer, P. J. (2017). Using event-related potentials to inform the neurocognitive processes underlying knowledge extension through memory integration. *Journal of Cognitive Neuroscience, 29*, 1932–1949.

Varga, N. L., Stewart, R. A., & Bauer, P. J. (2016). Integrating across episodes: Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. *Journal of Experimental Child Psychology, 145*, 48–63.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement, 24*(3), 4–12.